Landscape and Urban Planning xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Landscape and Urban Planning

journal homepage: www.elsevier.com/locate/landurbplan

Research Paper

A Bayesian approach to mapping the uncertainties of global urban lands

Zutao Ouyang[a,*], Peilei Fan[b], Jiquan Chen[c], Raffaele Lafortezza[d,a], Joseph P. Messina[c], Vincenzo Giannico[d], Ranjeet John[c]

[a] Center for Global Change and Earth Observations (CGCEO), Michigan State University, 1405 S. Harrison Road, East Lansing, MI 48823, USA
[b] School of Planning, Design, and Construction and Center for Global Change and Earth Observations (CGCEO), Michigan State University, East Lansing, MI 48823, USA
[c] Department of Geography, Environment, and Spatial Sciences and Center for Global Change and Earth Observations, Michigan State University, USA
[d] Department of Scienze Agro-Ambientali e Territoriali, University of Bari Aldo Moro, Via Amendola 165/A, 70126 Bari, Italy

ARTICLE INFO

Keywords:
Remote sensing
Bayesian
Urban
MODIS
Uncertainty
Hybrid

ABSTRACT

Global distribution of urban lands is one of the essential pieces of information necessary for urban planning. However, large disagreement exists among different products and the uncertainty remains difficult to quantify. We applied a Bayesian approach to map the uncertainties of global urban lands. We demonstrated the approach by producing a hybrid global urban land map that synthesized five different urban land maps in ca. 2000 at 1-km resolution. The resulting hybrid map is a posterior probability map with pixel values suggesting the probability of being urban land, which is validated by 30-m higher resolution references. We also quantified the minimum and maximum urban areas in 2000 for each country/continent based on subjective probability thresholds (i.e., 0.9 and 0.1) on our hybrid urban map. Globally, we estimated that the urban land area was between 377,000 and 533,000 km^2 in 2000. The credible interval of minimum/maximum urban area can help guide future studies in estimating urban areas. In addition to providing uncertainty information, the hybrid map also achieves higher accuracy than individual maps when it is converted into a binary urban/non-urban map using a probability threshold of 0.5. This new method has the ability to further integrate discrete site/location-based data, local, regional, and global urban land maps. As more data is sequentially integrated, the accuracy is expected to improve. Therefore, our hybrid map should not be regarded as a final product, but a new prior product for future synthesis and integration toward a "big data" solution.

1. Introduction

Although urban areas cover only a small percentage of the Earth's surface, their significance in altering climate, biogeochemistry, hydrology and other socioecological functions goes far beyond their boundaries (Fan, Xu, Yue, & Chen, 2017; Imhoff et al., 2004; Seto, Sánchez-Rodríguez, & Fragkias, 2010; Wu, 2014). Urban land use is also critical baseline information for many large-scale models (e.g., in developing future scenarios of land use, climate change, and species dispersal) and urban planning and management, but these models still lack accurate representation of urban areas (Oleson, Bonan, Feddema, & Vertenstein, 2008). Accurate delineations of the global distribution and uncertainty of urban area are therefore important (Alberti, 2005; Kaye, Groffman, Grimm, Baker, & Pouyat, 2006), especially for developing future landscape designs and policies, and for land surface modeling at regional and global scales.

Previous efforts in developing reliable global maps of urban areas within the context of other land cover types have been made by numerous international groups of academic and government sectors at relatively coarse resolutions (i.e., 300–1000 m). These products, however, rarely agree with each other (Fritz & See, 2008; Taubenbock, Esch, Felbier, Roth, & Dech, 2011). The disagreements also vary substantially in time and space (Fritz et al., 2011; Klotz, Kemper, Geiß, Esch, & Taubenböck, 2016) because of a lack in an united thematic definition of the land cover class and low overall accuracy among many products (Fritz & Seet, 2005; Liu, He, Zhou, & Wu, 2014). For example, there are at least 10 global-scale urban maps that can be used to identify urban area ca. 2000 at relatively coarse resolutions, but there is an order of magnitude difference in the estimation of total global urban areas among them (Potere, 2009). There are new products that are emerging at unprecedented spatial resolutions (e.g., 30–90 m) and enhanced reliability, such as the Global Urban Footprint (GUF) (Esch et al., 2013), the Global Human Settlement Layer (GHSL) (Pesaresi et al., 2013), and the Global Land Cover at 30 m resolution (GLC30).

* Corresponding author.
E-mail addresses: yangzuta@msu.edu (Z. Ouyang), fanpeile@msu.edu (P. Fan), jqchen@msu.edu (J. Chen), jpm@msu.edu (J.P. Messina), vincenzo.giannico@uniba.it (V. Giannico), ranjeetj@msu.edu (R. John).

https://doi.org/10.1016/j.landurbplan.2018.07.016
Received 19 September 2017; Received in revised form 23 March 2018; Accepted 21 July 2018

However, the same issues, i.e., significant differences in accuracy in different regions, still exist (Klotz et al., 2016).

Facing the above challenges, scientific investigations have been promoted to compare the quality and accuracy of different land cover maps and to quantify the spatial agreement and disagreement, aiming at hybrid maps through the integration of different products. Potere (2009) assessed and compared the quality and accuracy of eight global urban maps ca. 2000, and Klotz et al. (2016) cross-compared the accuracy of global settlement layers with the recently developed GUF and GHS product included. Fritz and See (2008) developed a disagreement map by comparing Global Land Cover 2000 (GLC-2000) (Fritz et al., 2003) and the MODIS land cover 1-km product (Friedl et al., 2002), which were further developed into uncertainty and spatial disagreement maps for user-specific applications (Fritz and See, 2008). A new global land cover map was created by Iwao et al. (2011) through the integration of GLC-2000, MODIS land cover, and the University of Maryland's 1-km Global Land cover product (Hansen, Defries, Townshend, & Sohlberg, 2000). A hybrid global land cover map was also constructed by See et al. (2015) using crowdsourcing validation data and geographically weighted regression – a method similar to that used to develop hybrid forest maps (Schepaschenko et al., 2015). The development of these hybrid products depends on large volumes of validation and training data acquired through crowdsourcing/volunteer–based programs, e.g., GeoWiki (Fritz et al., 2009) and the Degree Confluence Project (Iwao, Nishida, Kinoshita, & Yamagata, 2006).

Since many existing maps can also be used to describe urban land, a hybrid urban land map can be identified and delineated by combining the information from multiple sources to quantify the mapping uncertainty among different maps. Therefore, the goal of this research is to develop an innovative method to integrate different types of available open-access geospatial data into mapping uncertainties of urban land. We creatively applied Bays' rule to integrate multiple data and products, which is a replicable model that can integrate as many data as possible. To demonstrate this approach and its performance, we also developed the first hybrid global urban land map to describe the uncertainties of global urban land mapping efforts for 2000, using five global urban maps ca. 2000 (see Methodology section). The five global urban maps are sequentially integrated into an urban probability/uncertainty model by repeating the same model five times. The integration process to produce a hybrid map is meaningful under our premise that separately produced urban land maps could provide complementary information to our recognition of urban land, and thus improve urban land mapping if they were integrated and fused appropriately. A hybrid map like this will not only provide the spatial distribution of urban land, but more importantly, the reliability of each pixel as urban land and the magnitude of agreement among different products for users.

2. Methodology

2.1. Data

There are currently about 10 global products that provide portraits of urban land for different years. We decided to produce a hybrid map for 2000 because more global products representing the year 2000 exist than for other years. We included five global maps (Table 1): Global Land Cover 2000 v1.1. – GLC-2000 (Fritz et al., 2003), GlobCover v2 – GLOBC (Arino et al., 2007), Global Impervious Surface Area – IMPSA (Elvidge et al., 2007), MODIS Urban Land Cover 500 m – MOD500 (Schneider, Friedl, & Potere, 2009a), and MODIS Urban Land Cover 1 km-MOD1k (Friedl et al., 2002). There are other published maps related to urban extent, including the GRUMP urban extent (Center for International Earth Science Information Network – CIESIN – Columbia, IFPRI, Bank, & Centro Internacional de Agricultura-CIAT, 2011), the History Database of the Global Environment v3 (Goldewijk, 2005), and the Landscan ambient population (Bhaduri, Bright, Coleman, & Dobson, 2002). However, they are either at coarser resolution or have a very different thematic definition of urban lands than the five maps we used, and were thus excluded from this study. The new products, such as GHSL, GUF, and GLC30, were not included due to lack of detailed accuracy assessments (i.e., both user's and producer's accuracy) or availability for 2000. Another reason that we excluded them is because we aimed at a coarse resolution (1000 m) product and a high-resolution hybrid map can be produced in the future using our approach.

We define "urban lands" as places dominated by built-up environment (Ouyang, Fan, & Chen, 2016; Potere, 2009) that are represented by "artificial surfaces and associated areas" in GLC-2000 and GLOBC, and by the "urban and built-up" class in MOD500 and MOD1k. The IMPSA is a percentage map of impervious area. Consequently, we converted them into a binary urban land map with a threshold of 20% (Potere, 2009). All maps were resampled to 1-km resolution and then converted to the same geographical coordinates – WGS84.

There are overlaps of the input datasets and methodologies among the five products. For example, MODIS1k, GLOBC and IMPSA used nighttime lights (Potere, 2009), while MODIS1k and MODIS500 shared some common steps in their algorithms (Schneider et al., 2009a). However, each product has distinct features in terms of its classification approach, main datasets, and the working group involved. Therefore, we assume that they are independent, with each contributing some novel and independent information.

2.2. Study area and experimental setup

Our study area is global, covering 180W–180E and 65S–75N. This area was divided into 43,200 by 15,722 pixels at 1-km resolution. For each pixel, we estimated the uncertainty of it being urban land using probability, and we applied Bayes' rule of probability to make statistical inferences. Each of the five maps classified a pixel as urban land or non-urban land, which is analogous to a Bernoulli test with either a positive or negative result – a fundamental philosophy that the classification process is regarded as a test and the classification result is the binary test result. Because none of the classifications are perfect but with errors (i.e., omission and commission errors, which corresponds to false negative and false positive), we could not 100% trust any one, suggesting that a better approach is to combine information from all to make a more reliable inference. We applied Bayes' rule to sequentially update our statistical inference of the probability of a pixel being urban land.

2.3. Bayes' rule, hypothesis, and the prior

Bayes' rule is the mathematical method to update one's probabilities about a certain statement of the real world (i.e., a hypothesis.) In its simplest form, Bayes' rule is stated as:

$$Pr(H|Data) = (Pr(H) \times Pr(Data|H))/Pr(Data) \quad (1)$$

where *Data* stands for new data (e.g., test results), and *H* stands for any hypothesis whose probability may be affected by evidence from *Data*. *Pr* (*H*) is the prior probability of the hypothesis before *Data*, *Pr* (*Data*) is probability of observing *Data* under all possible hypothesis, and *Pr* (*Data*|*H*) is the likelihood (the probability of the observing *Data* under the hypothesis *H*) and *Pr* (*H*|*Data*) is the posterior probability after *Data* is observed. In other words, the posterior probability of a hypothesis is proportional to the prior probability times the likelihood.

Because we are interested in how likely it is that a pixel is urban land, we formed our hypothesis as: A pixel is either urban land (U+) or non-urban land (U−), and the probability of which follows a Bernoulli distribution:

$$\Pr(X) = \begin{cases} pif\ X = U+ \\ 1-pif\ X = U- \end{cases} \quad (2)$$

Table 1
The five global urban land maps used in this study. NOAA – National Oceanographic and Atmospheric Administration; NASA – National Aeronautics and Space Administration.

Name	Description	Source	Type	Resolution
GLC-2000	Global Land Cover 2000 v1.1	European Commission Joint Research Center (http://www.gvm.jrc.it/glc2000)	Thematic (22 classes)	1000 m
GLOBC	GlobCover V2	European Commission Joint Research Center (http://due.esrin.esa.int/page_globcover.php)	Thematic (22 classes)	300 m
IMPSA	Global Impervious Surface Area	US National Geophysical Data Center, NOAA) (https://ngdc.noaa.gov/eog/dmsp/download_global_isa.html)	Continuous (percent impervious area)	1000 m
MOD500	MODIS Urban Land Cover 500 m	University of Wisconsin, Boston University, NASA (http://nelson.wisc.edu/sage/data-and-models/datasets.php)	Binary	500 m
MOD1k	MODIS Urban Land Cover 1 km	Boston University, NASA (http://www-modis.bu.edu/landcover)	Binary	1000 m

where $Pr\,(x = U+) + Pr\,(x = U-) = 1$. Here we focused our inference on $Pr\,(x = U+) = p$. To simplify the mathematical symbol, we omit x and use $Pr\,(U+)$ for $Pr\,(x = U+)$ and $Pr\,(U-)$ for $Pr\,(x = U-)$. Prior to any tests to make the statistical inference about p, we need to quantify the prior probability of p. The prior p of a Bernoulli distribution is commonly interpreted as the prevalence in a large population (Cowles, 2013; Pritchard & Tebbs, 2011). Therefore, we used the prior estimation of the fraction of urban land as an informative prior for all pixels. This fraction, while has been estimated by various studies (Angel et al., 2016; Center for International Earth Science Information Network, 2013; Liu et al., 2014), varied between 0.1% and 10%. So far, there is no evidence that any one of the estimates is superior or more precise; we thus applied a uniform prior as:

$$p \;\; uniform\,(0.001,\, 0.1) \tag{3}$$

2.4. Inference of posterior probability

We used 1 to represent a positive test result from a map (i.e., when a pixel is classified as urban land) and 0 as a negative result (i.e., when it is classified as non-urban land), while the performance of the test was determined by the map-wide accuracies. Taking GLC-2000 as an example, if a pixel received a positive test result from GLC-2000, then:

$$Pr(U+|1) = Pr(U+) \times Pr(1|U+)/(Pr(1)) \tag{4}$$

$$Pr(1) = Pr(1|U+) \times Pr(U+) + Pr(1|U-) \times Pr(U-) \tag{5}$$

Similarly, if a pixel received a negative test result from GLC-2000, then:

$$Pr(U+|0) = Pr(U+) \times Pr(0|U+)/(Pr(0)) \tag{6}$$

$$Pr(0) = Pr(0|U+) \times Pr(U+) + Pr(0|U-) \times Pr(U-) \tag{7}$$

In Eqs. (4)–(7), $P\,(U+|1)$ and $P\,(U+|0)$ are posterior probabilities after knowing the test (classification) result from GLC-2000; $Pr\,(U+)$ is the prior probability of a pixel being urban land; $P\,(U-) = 1 - P\,(U+)$; and $P\,(1|U+)$, $P\,(0|U+)$, $P\,(1|U-)$, and $P\,(0|U-)$ are the likelihoods (conditional probabilities) that quantify the performance of the binary classification test. $P\,(1|U+)$ is the sensitivity (true positive rate) that measures the proportion of urban land that is correctly identified as urban land. $P\,(0|U-)$ is the specificity (true negative rate) that measures the proportion of non-urban land that is correctly identified as such. $P\,(0|U+)$ is the omission error rate (false negative rate), i.e., the proportion of urban land that is mistakenly identified as non-urban land, and $P\,(+|U-)$ is the commission error (false positive rate), i.e., the proportion of non-urban land that is mistakenly classified as urban land; also sensitivity + omission error = 1 and specificity + commission error = 1. These conditional probabilities were determined by map-wide pixel-level accuracy assessment. Potere (2009) assessed the map-wide pixel-level sensitivity and specificity of eight global urban maps (including the five products we used) using nearly 10,000 high-resolution GE-validation sites (Table 2). Therefore, we adopted their values of test sensitivity and specificity. Starting with GLC-2000, the values for P (1|U+), P (0|U+), P (1|U−), and P (0|U−) are 0.42, 0.58, 0.023, and 0.977, respectively (Table 2). Substituting these values into Eqs. (4)–(7), the posterior probability of $Pr\,(U+)$ can then be updated, which equals to P (U+|1) or $P\,(U+|0)$, depending on whether the pixel is classified as urban or non-urban on GLC-2000. We denote the posterior probability of $Pr\,(U+)$ for each pixel after updating with GLC-2000 as $Pr\,(U+|\text{GLC-2000})$. $Pr\,(U+|\text{GLC-2000})$ was then regarded as a new prior for incorporating another test (map) (e.g., GLOBC). The same procedure was repeated five times to sequentially incorporate all five tests (maps)- GLC-2000, GLOBC, IMPSA, MOD500, MOD1k, and our final inference of $Pr\,(U+)$ is the posterior probability: $Pr\,(U+|\text{GLC2000, GLOBC, IMPSA, MOD500, MOD1k})$ (Fig. 1).

Table 2
The conditional probabilities used to calculate the likelihood and normalization terms for Eqs. (4)–(7). The conditional probabilities are the sensitivity ($P\,(1|U+)$), omission error ($P\,(0|U+)$), specificity ($P\,(0|U-)$), and commission error (($P\,(1|U-)$). They were provided by Potere (2009) using nearly 10,000 high-resolution GE- sites.

Map	$P\,(1\|U+)$	$P\,(0\|U+)$	$P\,(1\|U-)$	$P\,(0\|U-)$
GLC-2000	0.42	0.58	0.013	0.987
GLOBC	0.39	0.61	0.009	0.991
IMPSA	0.60	0.40	0.011	0.989
MOD500	0.75	0.25	0.020	0.980
MOD1k	0.70	0.30	0.031	0.969

To implement our sequential update of $Pr\,(U+)$ for each pixel (Fig. 1), some assumptions are made. First, although neighboring pixels might be spatially auto-correlated, we treated them independently by ignoring the horizontal correlation in space. Therefore, separate sequential updating for each pixel was run independently through a loop. Second, for each pixel we assume the five vertical "observations" (i.e., classification results) from the five products are independent and exchangeable. Therefore, the order of the sequence in Fig. 1 is interchangeable mathematically and semantically and would not affect the final results. Third, the accuracy of each product might vary for different regions even though the same classification techniques were applied. However, we assume that the test reliability for all pixels (i.e. the conditional probabilities) are the same because accuracy assessments for different pixel groups are not available.

2.5. Validation of the hybrid map

Unlike the five global maps used for integration, our hybrid map resulting from sequential updating based on Bayes' rule is a fuzzy classification map, where pixel values represent the probability of being urban land (i.e., uncertainty of urban classification). We validated the hybrid uncertainty in two ways. National land cover data (NLCD) in 2001 for the contiguous US and mainland China at 30-m resolution in 2000 were regarded as our references. First, to confirm that the probability values are meaningful in representing mapping uncertainties, we randomly selected 2000 potential urban pixels (probability > 0.1)

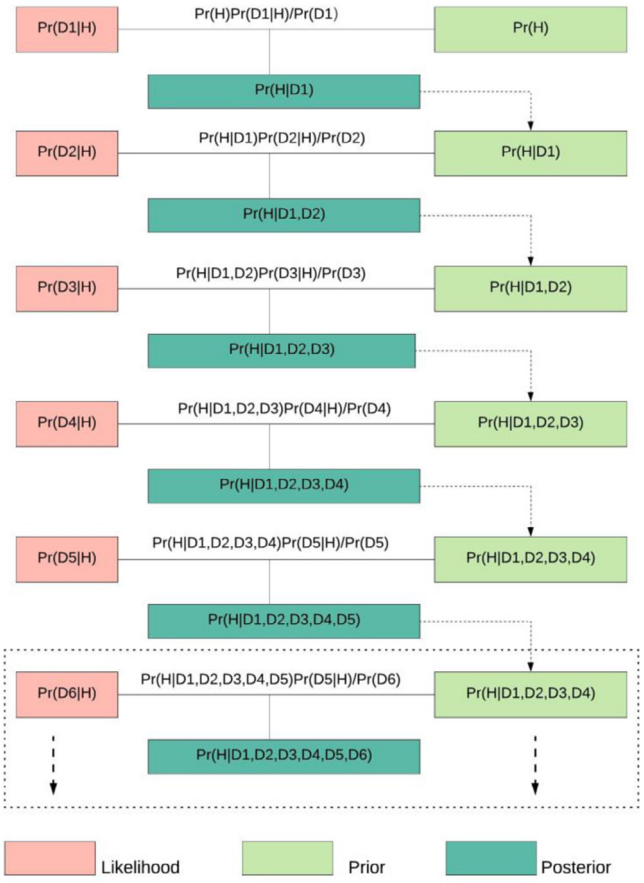

Fig. 1. Working flowchart demonstrating the sequential integration of five global urban maps through Bayesian sequential analysis. H is the hypothesis that a pixel is urban land, and D1, D2, D3, D4, D5 are the classification results from GLC-2000, GLOBC, IMPSA, MOD500, MOD1k, respectively. The classification results will be either urban land (denoted as 1 in Eqs. (4)–(7) or non-urban land (denoted as 0 in Eqs. (4)–(7). D6 represents any future dataset that will be available. This process chain was applied for each pixel separately.

from the hybrid map for the contiguous US to check if the low/high commission error corresponded with high/low probability values. Second, we checked if the hybrid map provided improved accuracy in addition to the uncertainty. Here, we first converted the hybrid probability maps into a binary urban vs. non-urban map using the threshold 0.5, and then compared it with the 30-m resolution reference data. We did this independently for the contiguous US and Mainland China, which represented different urban morphology from developed countries and developing countries, respectively. One hundred thousand 1×1 km squares were randomly selected for the contiguous US and Mainland China, respectively. The one hundred thousand pixels were then converted into square polygons. When > 50% of the area is covered by urban land class from 30-m NLCD, the polygon/pixel is referenced as urban land; otherwise it is classified as non-urban land. The reference data were compared with our categorized binary hybrid map and the MOD500 for computing both producers' and users' accuracy for comparison. In this study, we only compared with MOD500 because it is the most accurate product, as compared to the other four maps (Potere, 2009; Schneider, Friedl, & Potere, 2009b).

2.6. Analysis of the hybrid map

The hybrid map provides the probabilities of being urban land for each pixel, which is our best estimate based on our previous knowledge and the data from the five global urban products. However, we are also interested in region-based uncertainties. Consequently, we categorized urban probabilities into classes and calculated the total area of different classes. Extremely low/high values suggest very high certainty in whether they are non-urban (e.g., < 0.1) or urban (e.g., > 0.9), while the medium values (0.1–0.9) indicate that the degree of uncertainty lies in a medium range for which we cannot make a decision. Based on our subjective beliefs, thresholds can be applied to these probabilities to convert urban classification probabilities/uncertainties into binary urban classification and create credible intervals of the urban areas. Similar to the frequentists' use of 0.05/0.1 and 0.95/0.9 as probability thresholds for significance tests, we chose 0.1 and 0.9 as the two "magic" probability thresholds. If the chance of a pixel being urban is < 0.1, we believe that this chance is so low that it must be non-urban. In contrast, if the chance of a pixel being urban land is as high as 0.9 or higher, we believe this chance is so high only because it is urban land. For other values, the pixel may or may not be urban land, but higher values suggest a higher probability of a pixel being urban land. Accordingly, we define a minimum urban area of a country as the total area of all pixels with a > 0.9 probability of being urban land, and a maximum urban area as the total area of all pixels with a probability value > 0.1. The two numbers create a credible interval of the total urban land area for a region. These minimum and maximum urban areas were calculated for each country and continent using spatial statistics. The interval span, i.e., the total area of all pixels with a probability of being urban land > 0.1 minus the area of all pixels with a probability of being urban land > 0.9, indicates the total area of uncertain classification, which was calculated for each country to generate meaningful comparisons. We grouped all countries into the Organization for Economic Co-operation and Development (OECD) and non-OECD countries, as developed/high income and developing/low income countries may have different urban structures/morphologies. Although there are numerous lists of developed/developing countries and high income/low income countries, we utilized the OECD members to represent developed/high income countries, while regarding other non-OECD members as developing/low income countries.

3. Results

3.1. Validation and comparison

We confirmed that the probability value on our hybrid map is meaningful in representing uncertainty in urban land mapping, with higher values representing higher mapping accuracy (i.e., high user accuracy or low commission errors) (Table 3). By assigning a probability of > 0.1 as urban land class for the hybrid map, we found that the users' accuracy (0.952) is highest for pixels with a probability of > 0.9, whereas the users' accuracy for pixels with a probability between 0.5 and 0.9 declined significantly (0.764), and the user's accuracy for probability between 0.1 and 0.5 further declined (0.414). Our comparisons with MOD500 – the most accurate map among the five individual maps – in the contiguous United States and mainland China also confirmed that the sequential analysis based on Bayes' rule appeared reliable to produce an improved map with both higher users' and producers' accuracy of urban land classification (Table 4).

Table 3
The accuracy assessed for urban land (probability > 0.1) among different probability groups for the contiguous United States. Higher probability of the urban land class is associated with higher users' accuracy.

	Probability Group	User's accuracy (commission error)
Urban Land	≥0.9	0.951 (0.043)
	0.50–0.9	0.768 (0.236)
	0.1–0.5	0.424 (0.581)

Table 4
The accuracies between the hybrid map and MOD500 for the contiguous United States and Mainland China. A threshold of 0.5 was used to convert the hybrid probability map into binary urban land class (> 0.5) and non-urban class (≤0.5). The users' accuracy and producers' accuracy are reported using 100, 000 validation 1 × 1 km polygons in all cases.

	Hybrid Map		MOD500	
	User's accuracy	Producer's accuracy	User's accuracy	Producer's accuracy
United States	0.679	0.757	0.635	0.742
China	0.375	0.543	0.264	0.518

3.2. The hybrid urban map

Our hybrid urban map is an uncertainty map, with the pixel value suggesting the probability of being urban land. Due to the small proportion of urban areas, it is difficult to visually see the details at the global scale with 1-km resolution. We therefore resampled our hybrid map to 60-km resolution (Fig. 2). Additionally, six city areas, selected from developed and developing countries, were zoomed in at the original 1-km resolution for visualizing the details (Fig. 2). As expected, it is discernible that high probability was mainly observed in city core areas, while low and medium probabilities were found in rural–urban fringes. This indicates the easiness of classification of urban core areas and high agreement among the five urban maps. There is a decreased agreement among different urban maps toward the periphery, which shows an increased difficulty of classification. Additionally, it appeared that more low- and medium-probability pixels were around the cities in the developing countries (e.g., Guangzhou, Bangkok, and New Delhi) than those in the developed countries (e.g., London, New York, and Tokyo). The global maps at 0.5-degree resolution reveals a global pattern of urbanization (Fig. 2). The coastal areas of the United States, China, and India, the Great Lakes region, the North China Plain, Western Europe, and Japan were the most urbanized areas (Fig. 2).

3.3. Regional urban areas and uncertainties

The minimum and maximum of total urban area was estimated for each continent (Fig. 3) and country (Table S1). We assumed that the true total urban area of a region must be between the minimum and maximum area estimated from our hybrid map (Table 2). The minimum total urban area of our estimates is larger than the smallest total urban areas among the five individual maps, while the maximum total urban area is smaller than the largest total urban areas estimated from five previous maps (Fig. 3). This phenomenon reflected the "shrinking" effect of Bayesian analysis, where the posterior estimate is the result of shrinking the individual-based data toward a center value determined by all data and priors. As a result, the extreme estimates can be eliminated to increase the confident level of estimating regional total urban areas (Fig. 3). Globally, we estimated that the urban land area is between 377,000 and 533,000 km^2.

The total area of pixels with a probability of 0.1–0.9 was the quantity for the uncertain classification in a country (Table S1). China and the US are the two countries with the most uncertain urban land

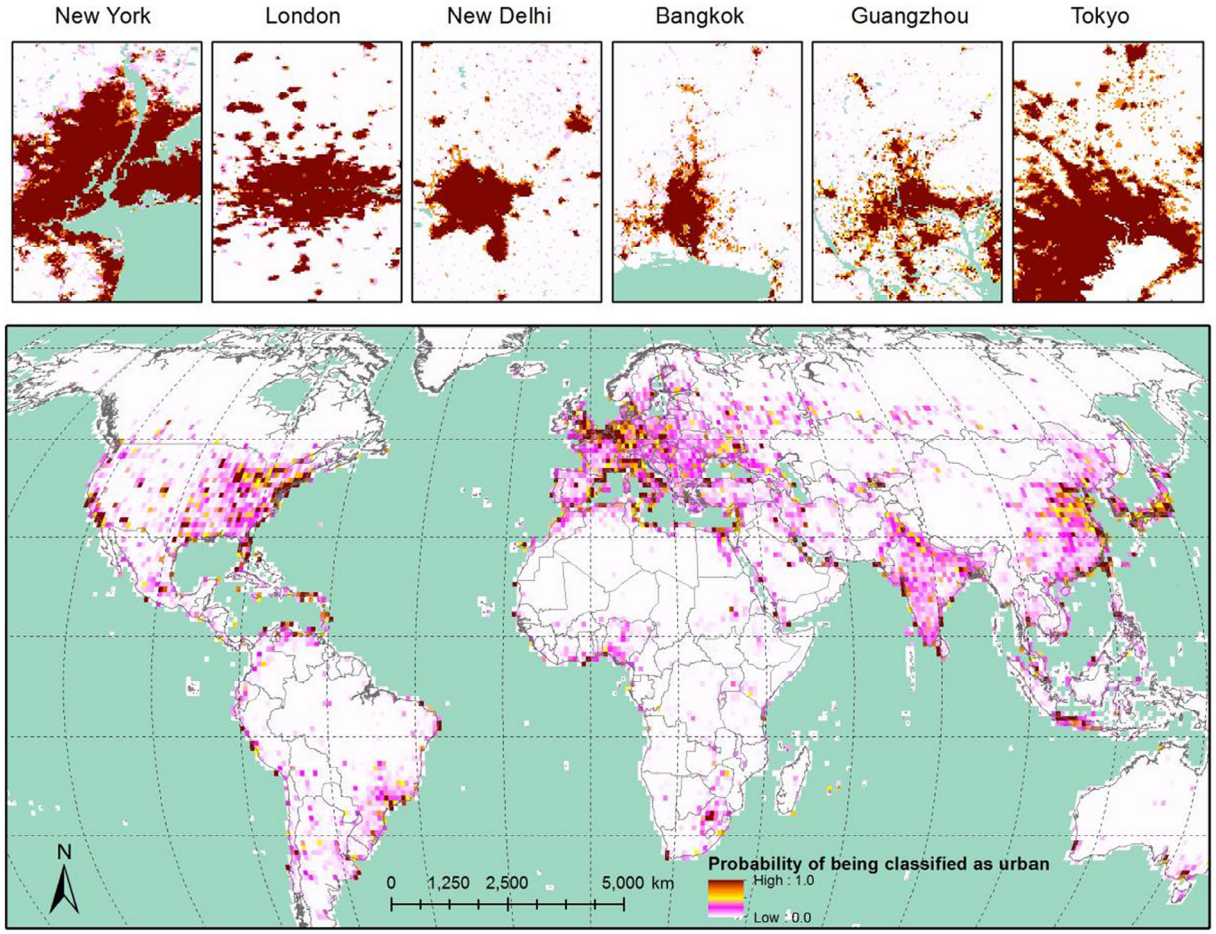

Fig. 2. Hybrid global urban maps resampled at 60-km resolution produced by sequential inference based on Bayes' rule. Six cities (New York, London, New Delhi, Bangkok, Guangzhou, and Tokyo) were enlarged from their original 1-km resolution to show the detail at a 1:2,500,000 scale.

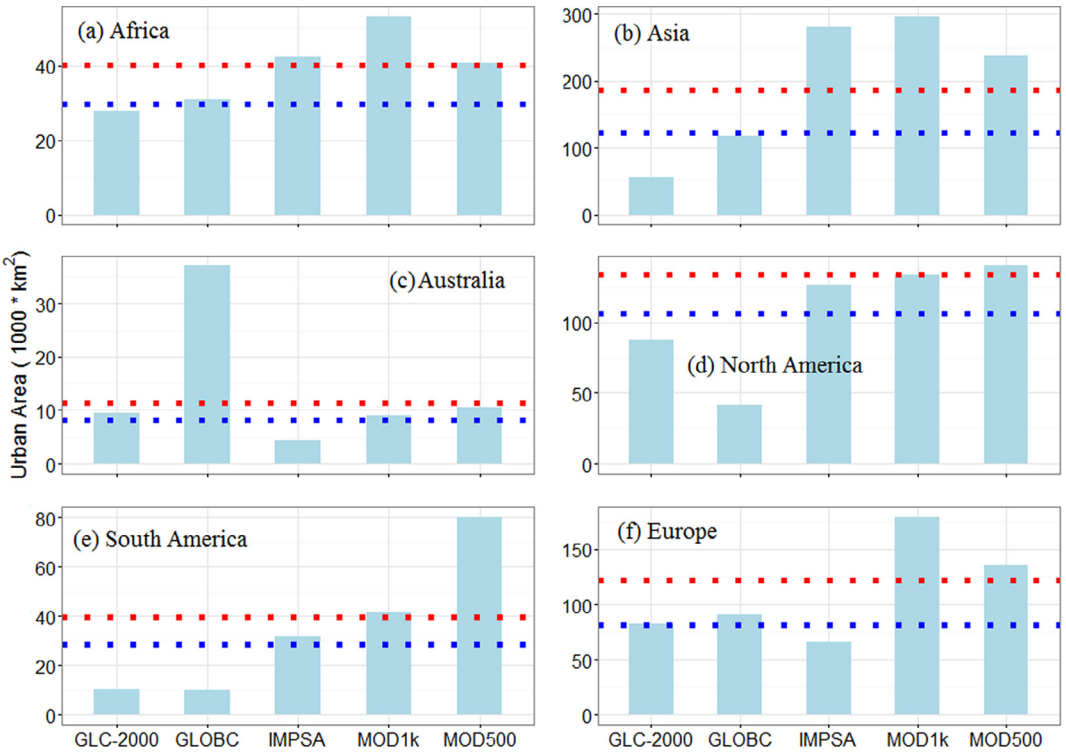

Fig. 3. Continental extents of urban lands (thousand km^2). The dotted blue line is an estimate of urban area based on the hybrid map with a threshold probability of 0.9 (minimum area) and the dotted red line is an estimate of urban area based on the hybrid map with a threshold probability of 0.1 (maximum area). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

areas, which represent the two largest economies in the OECD and non-OECD countries, respectively. For the non-OECD countries, Russia and India rank as the second and the third after China, while for the OECD members, Germany, Italy, and Japan followed the US in total amount of uncertain areas.

4. Discussions

4.1. Interpretation of the results

Our results have several implications. First, to echo our hypothesis, the resulting probability is interpreted as how likely it is that a pixel is urban land. This interpretation is validated to be meaningful, as it appeared that pixels with higher probability values were also more likely to be urban in reality (i.e., higher user's accuracy, Table 3). Therefore, the probabilities of being urban land in the urban–rural transitional areas are lower than the city core areas (Fig. 1). However, the probability on the hybrid map is a statistical inference of the real probability of being urban. It will not approach the real probability of being urban until a large number of products are integrated. Currently, it represents our best guess from the available information (i.e., 5 products). With more (massive) data being integrated into our modeling approach, future delineations of urban areas will be further improved.

Secondly, the probability value of pixels are also in accordance with the probability of being classified as urban, as when a pixel is very likely urban land, it also has a high chance of being detected as urban land by various maps/techniques. This the rationale for us to use the conditional probabilities (i.e., $Pr\,(1|U+)$, $Pr\,(0|U+)$, $Pr\,(1|U-)$ and $Pr\,(1|U-)$) from maps to inference the real probability of a pixel being urban land. Consequently, extreme values at both ends of probability distribution (i.e. > 0.9 or < 0.1) would suggest high levels of agreement among the five maps and high confidence with which we can safely classify it as urban land or non-urban land. When five maps all classified a pixel as urban land, it would lead to an extremely high probability value of the pixel on the hybrid map; similarly, when five maps all classified a pixel as non-urban land, it would lead to an extreme low probability value. Future attention should be paid to medium-scale values (i.e., 0.1–0.9) that represent a lower magnitude of agreement among the five products. Compared to other studies that have quantified the disagreements/agreements between only two of the products (Fritz and See, 2008; Fritz et al., 2009, 2011; See et al., 2015), our approach was able to include an unlimited number of products through sequential analysis (Fig. 2).

We achieved greater improved accuracy for urban land mapping than any input maps. Although MOD500 is considered the most accurate coarse resolution global maps (Table 2), the resulting hybrid map through the integration of other maps had higher accuracy than MOD500 for the contiguous United States and Mainland China (Table 4) when categorizing urban land/non-urban land using a probability threshold of 0.5. This echoes our premise that none of the individual maps was superior for all places, while the low-accuracy data can still contribute important mapping information for improvement. The increased users' accuracy of the hybrid map compared to that of MOD500 suggests that some mistakenly detected urban land pixels by MOD500 may not be confirmed by the other four products, and were therefore excluded from the hybrid map. The increased producers' accuracy further suggests that some mistakenly omitted urban land pixels were re-identified by the other four products. In sum, the improved accuracy is attributed to the use of multiple datasets.

4.2. Uncertain urban areas

Urban population and gross domestic product (GDP) growth have been widely regarded as important drivers or indicators of urban expansion (Ouyang et al., 2016; Zhang & Seto, 2011). However, it is not clear whether the uncertainty in mapping urban land is also correlated with these socioeconomic conditions. With a large number of available classifications (i.e., different maps), the chance for correctly

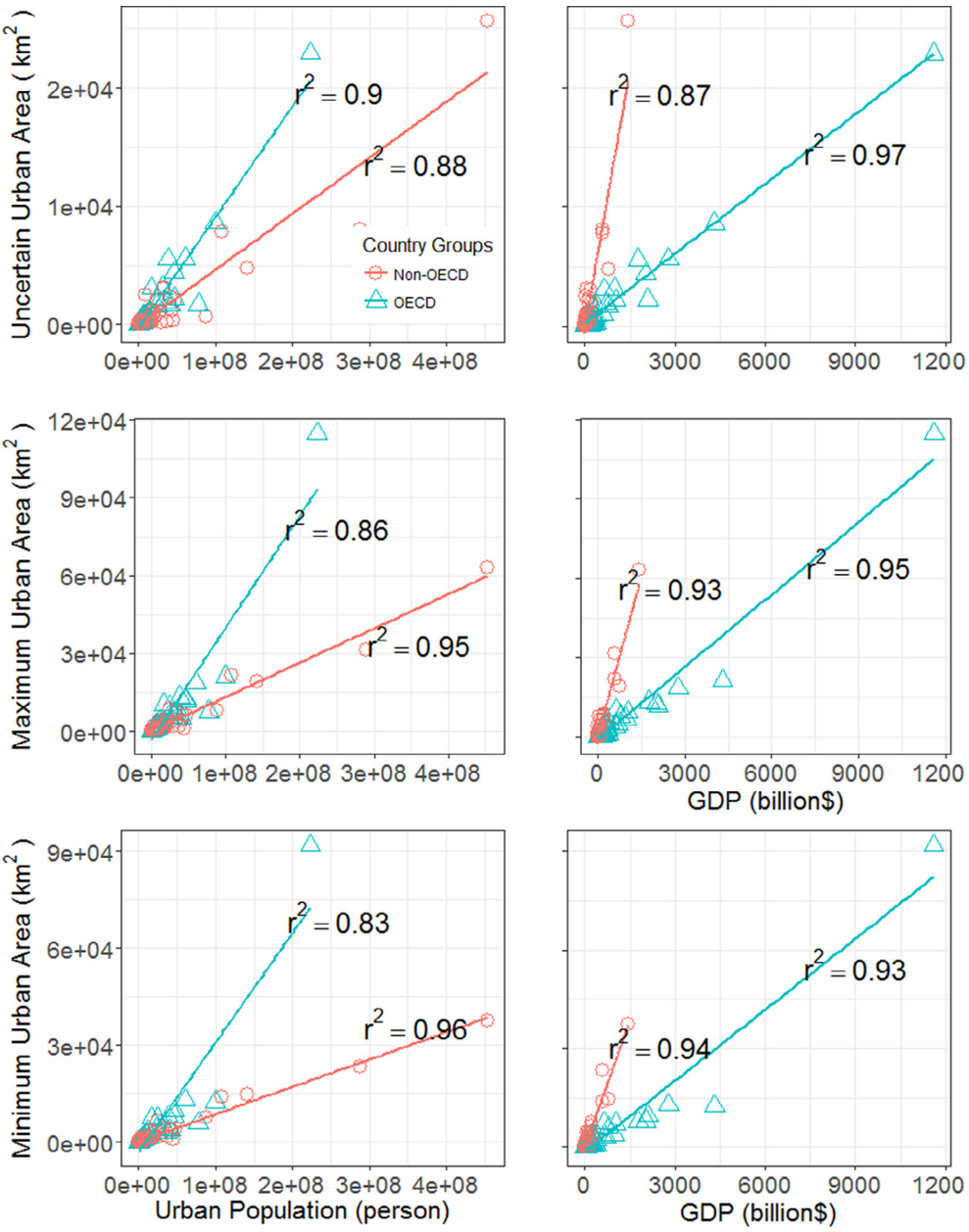

Fig. 4. The relationship between uncertain urban areas, minimum urban areas, and maximum urban areas with urban populations (a, b and c) and GDP (d, e, and f).

recognizing urban land could be affected by the urban landscape settings. In other words, there is a higher probability of being classified as urban land when an area has more urban features (e.g., more impervious surface, less bare soil and vegetation). Conversely, if an area has fewer urban features, it would more difficult to recognize and less accurately identified using any technique. While GDP and urban population can reflect urbanization levels and intensification, it is of interest to investigate their relationships with the uncertainty in mapping urban land. Using the secondary statistics GDP and urban population from the World Bank (http://www.worldbank.org/), we found that they are highly associated with the maximum and minimum total urban land area (quantified by the probability thresholds of 0.1 and 0.9, respectively). We found an interesting, different linear relationship for OECD members (developed/high income countries) and non-OECD countries (developing/low income countries) (Fig. 4). For the OECD countries, urban area per unit supports fewer people but produces higher economic outcomes (Fig. 4), which suggests a potential difference in urban structures/morphologies between the two groups. The total area of uncertain urban land is also highly correlated with GDP and urban population, suggesting co-linearity between the total uncertain urban area and the total urban area. Altogether, these relationships demonstrate that the mapping uncertainty is potentially related to the magnitude of urbanization and the status/structure of urbanization across the space. Uncertain urban land areas are commonly found in urban–rural fringes, small cities or towns with mixed pixels that has a high percentage of bare soils and vegetation cover (Cockx, Van de Voorde, & Canters, 2014; Irwin and Bockstael, 2007; Mertes, Schneider, Sulla-Menashe, Tatem, & Tan, 2015). More industrialized and developed countries tend to have larger urban areas and agglomeration (Decker, Kerkhoff, Moses, Kuhnert, & West, 2007; Mills & Peng Tan, 1980), which may be associated with more urban–rural transitional areas that are more difficult to detect. Under the same level of economic productivity, there is also divergence of the uncertainty in classifying urban land between developing/low income and developed/high-income countries (Fig. 4). This may be due to the fact that urban areas in developing countries have more bare soil, irregular

vegetable cover, unregulated slums, and broken water surfaces that create mixed pixels where spectra confused with impervious areas (Lu & Weng, 2006; Zhang, Chen, & Lu, 2015). Finally, data with high spatial resolutions, more ground surveys, and other ancillary data and technological sources in developed/high income countries may be richer for classification, which would in turn increase classification accuracy.

4.3. Limitations and future research

In an era of "big data", Bayesian modeling has received a lot of attention in processing spatial data, such as analyzing species distribution (Latimer, Wu, Gelfand, & Silander, 2006) and diversities (Latimer et al., 2006), modeling spatiotemporal variation in dynamic models (Stroud, Muller, & Sanso, 2001), and estimating urban growth potential (Smith et al., 2017). Herein, we extended the application of Bayes' rule to land use/land cover mapping and integration. This method, as we demonstrated in Fig. 1, can evolve naturally into a "big data" solution when additional new data become available. Although we were only able to acquire five products for incorporation in this study, other maps (e.g., Zhou et al., 2015), when available with the necessary information and access, can be easily included to update and improve our results with the method described here. Accordingly, our result is not an end product but rather the adaptive new prior for future data fusion and integration. The philosophy of "big data" solutions suggests that simple algorithms with more data will work better than more complex algorithms with less data. It is therefore ideal to incorporate as many datasets as possible into the sequence to improve the outcomes.

Although we only used maps that have global coverage in this study, the process would also work with data of varying spatial extents. Regional, national and provincial maps, and even single pixel/site data could be further integrated to the Bayesian sequential updating we proposed here. Because we treated each pixel separately, it is only necessary to modify the loop (Fig. 1) to exclude pixels outside of that region when updating with a map/data that only covers a proportion of the globe. Many countries and cities have highly accurate high-spatial resolution urban maps (e.g., Başnou et al., 2013; Fan, Chen, & John, 2016; Homer et al., 2007), which can be also integrated into our Bayesian sequence to improve the quality of the global map, assuming an accuracy assessment is reported. We did not perform these exercises due to lack of access to these data. Instead, we used global coverage data to demonstrate a generalizable and applicable methodology and philosophy. High-resolution global urban land products (e.g., GHSL, GUF, GLC30, etc.), especially for the periods after 2010, are emerging at unprecedented speeds due to advancements in high-resolution data sources, cloud storage capacity, and hyper performance computation. Our approach can be readily adopted to produce a high-resolution hybrid map for any period in the future where there are sufficient high-resolution urban land maps. Our coarse resolution hybrid map can be resampled to a higher resolution, and used as a prior layer for future data integration through a Bayesian chain.

Moreover, this "big data" age makes it easy for the public to be involved in scientific research. Many ongoing projects, such as geo-wiki, provide platforms and environments where both ordinary citizens and professionals can volunteer observations (visual classification) of predefined pixels. Inclusions of these projects into land use land cover mapping may emerge in the near future. When such observations are crowd-sourced, it will provide additional large data that can be integrated into a global map, or used separately to validate global maps. However, to apply a Bayesian framework for sequential updates, it is necessary to have an independent uncertainty assessment of the volunteered geographic information on land cover (Comber et al., 2013). Currently, such data are insufficient. Although we did not include such an analysis, it is an important direction for future investigations.

One aspect to improve our Bayesian inference is to use region-based classification accuracy (i.e., test reliabilities). We have used the globally based classification accuracy for all pixels from the same product, while several studies had suggested classification accuracy varied across regions even with the same classification procedure (i.e., the same test instrument). This is a widely spread issue that the same test carries different false positive and false negative rates when it is applied to different population. Therefore, using more precise test reliability values for pixels in different regions may potentially improve the inference of posterior probability. Another limitation of our current model is that it only applies to binary classes, which is good for mapping forest vs. non-forest, water vs. non-water, cropland vs. non-cropland, but not for multiclass classifications. We also only selected products that were developed with similar thematic definitions of urban land, and thus minimized the issues of thematic accuracy (Fritz et al., 2011) and application-specific uncertainty. However, when extending our methods to include multiple classes and multiple maps with inconsistent thematic definitions, additional cautions are needed. To extend this inference of the binary data parameter *Pr* (p) to multiple land use types, we can expand it by applying on a Multinomial distribution (e.g., Dirichlet distribution) instead of a Bernoulli distribution.

5. Conclusions

In this study, we proposed and applied Bayes' rule to sequentially update the probability of a pixel being urban land through integrating multiple existing maps, and illustrated its implementation by producing the first hybrid global urban map for the year 2000. This map provides the probability of being urban land for each 1×1 km pixel, and is an integrated product based on five previously employed global urban maps (i.e., MOD500, MODIS1k, GLOBC, IMPSA, and GCG-2000). The hybrid map revealed regions and places where more studies are needed to improve urban land mapping for planning and management. We also quantified the minimum and maximum urban areas in the year 2000 for each country/continent based on our hybrid urban map. We estimated the global urban land area to be between 377,000 and 533,000 km^2. We found that mapping uncertainty was related to economic and urbanization level, with distinct differences between high income/developed and low income/developing countries. This implies that disagreement among different global urban maps and mapping uncertainty may be related to the complexity of the urban systems rather than only the differences in mapping methods and data sources. Finally, this paper contributes a "big data" solution to land use/land cover mapping as various sources of data (observations, classifications) can be exploited by this simple algorithm (i.e., Bayes' rule) to increase the mapping quality of urban land.

Acknowledgement

We would like to acknowledge the financial support from the National Aeronautics and Space Administration's (NASA) Land Cover and Land Use Change Program (LCLUC) through its grant to Michigan State University (Grant No. NNX15AD51G). We thank Damien Sulla-Menashe for providing us the MOD1k data. We thank Connor Crank for editing the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.landurbplan.2018.07.016.

References

Alberti, M. (2005). The effects of urban patterns on ecosystem function. *International Regional Science Review, 28*(2), 168–192. https://doi.org/10.1177/0160017605275160.

Angel, S., Parent, J., Civco, D. L., & Blei, A. M. (2016). Atlas of urban expansion – the

2016 edition, Volume 1: Areas and Densities. Choice Reviews Online (Vol. 50). http://doi.org/10.5860/CHOICE.50-1227.

Arino, O., Gross, D., Ranera, F., Leroy, M., Bicheron, P., Brockman, C., ... Weber, J.-L. (2007). GlobCover: ESA service for global land cover from MERIS. *2007 IEEE International Geoscience and Remote Sensing Symposium* (pp. 2412–2415). IEEE.

Başnou, C., Álvarez, E., Bagaria, G., Guardiola, M., Isern, R., Vicente, P., & Pino, J. (2013). Spatial patterns of land use changes across a mediterranean metropolitan landscape: Implications for biodiversity management. *Environmental Management, 52*(4), 971–980. https://doi.org/10.1007/s00267-013-0150-5.

Bhaduri, B. L., Bright, E. A., Coleman, P. R., & Dobson, J. E. (2002). LandScan: Locating people is what matters. *Geoinformatics, 5*(2), 34–37.

Center for International Earth Science Information Network - CIESIN - Columbia, IFPRI, I. F. P. R. I., Bank, T. W., & Centro Internacional de Agricultura-CIAT (2011). *Global Rural-Urban Mapping Project, Version 1 (GRUMPv1): Urban Extents Grid.* Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC).

Center for International Earth Science Information Network (2013). *Low Elevation Coastal Zone (LECZ) Urban-Rural Population and Land Area Estimates, Version 2.* Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC).

Cockx, K., Van de Voorde, T., & Canters, F. (2014). Quantifying uncertainty in remote sensing-based urban land-use mapping. *International Journal of Applied Earth Observation and Geoinformation, 31*, 154–166. https://doi.org/10.1016/j.jag.2014.03.016.

Comber, A., See, L., Fritz, S., Van der Velde, M., Perger, C., & Foody, G. (2013). Using control data to determine the reliability of volunteered geographic information about land cover. *International Journal of Applied Earth Observation and Geoinformation, 23*, 37–48. https://doi.org/10.1016/j.jag.2012.11.002.

Cowles, M. K. (2013). *Applied Bayesian statistics: With R and OpenBUGS examples.* Springer.

Decker, E. H., Kerkhoff, A. J., Moses, M. E., Kuhnert, C., & West, G. (2007). Global patterns of city size distributions and their fundamental drivers. *PLoS One, 2*(9), e934. https://doi.org/10.1371/journal.pone.0000934.

Elvidge, C. D., Tuttle, B. T., Sutton, P. C., Baugh, K. E., Howard, A. T., Milesi, C., ... Nemani, R. (2007). Global distribution and density of constructed impervious surfaces. *Sensors, 7*(9), 1962–1979. https://doi.org/10.3390/s7091962.

Esch, T., Marconcini, M., Felbier, A., Roth, A., Heldens, W., Huber, M., ... Dech, S. (2013). Urban Footprint Processor—Fully Automated Processing Chain Generating Settlement Masks From Global Data of the TanDEM-X Mission. *IEEE Geoscience and Remote Sensing Letters, 10*(6), 1617–1621. https://doi.org/10.1109/LGRS.2013.2272953.

Fan, P., Chen, J., & John, R. (2016). Urbanization and environmental change during the economic transition on the Mongolian Plateau: Hohhot and Ulaanbaatar. *Environmental Research, 144*, 96–112. https://doi.org/10.1016/j.envres.2015.09.020.

Fan, P., Xu, L., Yue, W., & Chen, J. (2017). Accessibility of public urban green space in an urban periphery: The case of Shanghai. *Landscape and Urban Planning, 165*, 177–192. https://doi.org/10.1016/j.landurbplan.2016.11.007.

Friedl, M., McIver, D., Hodges, J. C., Zhang, X., Muchoney, D., Strahler, A., ... Schaaf, C. (2002). Global land cover mapping from MODIS: Algorithms and early results. *Remote Sensing of Environment, 83*(1), 287–302. https://doi.org/10.1016/S0034-4257(02)00078-0.

Fritz, S., Bartholomé, E., Belward, A., Hartley, A., Stibig, H. J., Eva, H., ... Defourny, P. (2003). *Harmonisation, mosaicing and production of the Global Land Cover 2000 database (beta version).* Luxembourg: Office for Official Publications of the European Communities.

Fritz, S., McCallum, I., Schill, C., Perger, C., Grillmayer, R., Achard, F., ... Obersteiner, M. (2009). Geo-Wiki.Org: The use of crowdsourcing to improve global land cover. *Remote Sensing, 1*(3), 345–354. https://doi.org/10.3390/rs1030345.

Fritz, S., & See, L. (2008). Identifying and quantifying uncertainty and spatial disagreement in the comparison of Global Land Cover for different applications. *Global Change Biology, 14*(5), 1057–1075. https://doi.org/10.1111/j.1365-2486.2007.01519.x.

Fritz, S., See, L., McCallum, I., Schill, C., Obersteiner, M., van der Velde, M., ... Achard, F. (2011). Highlighting continued uncertainty in global land cover maps for the user community. *Environmental Research Letters, 6*(4), 44005. https://doi.org/10.1088/1748-9326/6/4/044005.

Fritz, S., & Seet, L. (2005). Comparison of land cover maps using fuzzy agreement. *International Journal of Geographical Information Science, 19*(7), 787–807. https://doi.org/10.1080/13658810500072020.

Goldewijk, K. K. (2005). Three centuries of global population growth: A spatial referenced population (density) database for 1700?2000. *Population and Environment, 26*(4), 343–367. https://doi.org/10.1007/s11111-005-3346-7.

Hansen, M. C., Defries, R. S., Townshend, J. R. G., & Sohlberg, R. (2000). Global land cover classification at 1 km spatial resolution using a classii cation tree approach. *International Journal of Remote Sensing, 21*(7), 1331–1364.

Homer, C., Dewitz, J., Fry, J., Coan, M., Hossain, N., Larson, C., ... Wickham, J. (2007). Completion of the 2001 national land cover database for the conterminous United States. *Photogrammetric Engineering and Remote Sensing ·, 70*(7), 337–341.

Imhoff, M. L., Bounoua, L., DeFries, R., Lawrence, W. T., Stutzer, D., Tucker, C. J., & Ricketts, T. (2004). The consequences of urban land transformation on net primary productivity in the United States. *Remote Sensing of Environment, 89*(4), 434–443. https://doi.org/10.1016/j.rse.2003.10.015.

Irwin, E. G., & Bockstael, N. E. (2007). The evolution of urban sprawl: Evidence of spatial heterogeneity and increasing land fragmentation. *Proceedings of the National Academy of Sciences of the United States of America, 104*(52), 20672–20677. https://doi.org/10.1073/pnas.0705527105.

Iwao, K., Nasahara, K. N., Kinoshita, T., Yamagata, Y., Patton, D., & Tsuchida, S. (2011). Creation of new global land cover map with map integration. *Journal of Geographic Information System, 3*(2), 160–165. https://doi.org/10.4236/jgis.2011.32013.

Iwao, K., Nishida, K., Kinoshita, T., & Yamagata, Y. (2006). Validating land cover maps with Degree Confluence Project information. *Geophysical Research Letters, 33*(23), L23404. https://doi.org/10.1029/2006GL027768.

Kaye, J. P., Groffman, P. M., Grimm, N. B., Baker, L. A., & Pouyat, R. V. (2006). A distinct urban biogeochemistry? *Trends in Ecology & Evolution, 21*(4), 192–199. https://doi.org/10.1016/j.tree.2005.12.006.

Klotz, M., Kemper, T., Geiß, C., Esch, T., & Taubenböck, H. (2016). How good is the map? A multi-scale cross-comparison framework for global settlement layers: Evidence from Central Europe. *Remote Sensing of Environment, 178*, 191–212. https://doi.org/10.1016/j.rse.2016.03.001.

Latimer, A. M., Wu, S., Gelfand, A. E., & Silander, J. A., Jr. (2006). Building statistical models to analyze species distributions. *Ecological Applications, 16*(1), 33–50. https://doi.org/10.1890/04-0609.

Liu, Z., He, C., Zhou, Y., & Wu, J. (2014). How much of the world's land has been urbanized, really? A hierarchical framework for avoiding confusion. *Landscape Ecology, 29*(5), 763–771. https://doi.org/10.1007/s10980-014-0034-y.

Lu, D., & Weng, Q. (2006). Use of impervious surface in urban land-use classification. *Remote Sensing of Environment, 102*(1), 146–160. https://doi.org/10.1016/j.rse.2006.02.010.

Mertes, C. M., Schneider, A., Sulla-Menashe, D., Tatem, A. J., & Tan, B. (2015). Detecting change in urban areas at continental scales with MODIS data. *Remote Sensing of Environment, 158*, 331–347. https://doi.org/10.1016/j.rse.2014.09.023.

Mills, E. S., & Peng Tan, J. (1980). A Comparison of urban population density functions in developed and developing countries. *Urban Studies, 17*(1), 3–321. http://journals.sagepub.com/doi/pdf/10.1080/00420988020080621.

Oleson, K. W., Bonan, G. B., Feddema, J., & Vertenstein, M. (2008). An urban parameterization for a Global Climate Model. Part II: Sensitivity to input parameters and the simulated urban heat island in offline simulations. *Journal of Applied Meteorology and Climatology, 47*(4), 1061–1076. https://doi.org/10.1175/2007JAMC1598.1.

Ouyang, Z., Fan, P., & Chen, J. (2016). Urban built-up areas in transitional economies of Southeast Asia: Spatial extent and dynamics. *Remote Sensing, 8*(10), 819. https://doi.org/10.3390/rs8100819.

Pesaresi, M., Huadong, G., Blaes, X., Ehrlich, D., Ferri, S., Gueguen, L., ... Zanchetta, L. (2013). A global human settlement layer from optical HR/VHR RS data: Concept and first results. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 6*(5), 2102–2131. https://doi.org/10.1109/JSTARS.2013.2271445.

Potere, D. (2009). *Mapping the world's cities: An examination of global urban maps and their implications for conservation planning.* Princeton University.

Pritchard, N. A., & Tebbs, J. M. (2011). Bayesian inference for disease prevalence using negative binomial group testing. *Biometrical Journal. Biometrische Zeitschrift, 53*(1), 40–56. https://doi.org/10.1002/bimj.201000148.

Schepaschenko, D., See, L., Lesiv, M., McCallum, I., Fritz, S., Salk, C., ... Ontikov, P. (2015). Development of a global hybrid forest mask through the synergy of remote sensing, crowdsourcing and FAO statistics. *Remote Sensing of Environment, 162*, 208–220. https://doi.org/10.1016/j.rse.2015.02.011.

Schneider, A., Friedl, M. A., & Potere, D. (2009a). A new map of global urban extent from MODIS satellite data. *Environmental Research Letters, 4*(4), 44003. https://doi.org/10.1088/1748-9326/4/4/044003.

Schneider, A., Friedl, M. A., & Potere, D. (2009b). A new map of global urban extent from MODIS satellite data. *Environmental Research Letters, 4*(4), https://doi.org/10.1088/1748-9326/4/4/044003.

See, L., Schepaschenko, D., Lesiv, M., McCallum, I., Fritz, S., Comber, A., ... Obersteiner, M. (2015). Building a hybrid land cover map with crowdsourcing and geographically weighted regression. *ISPRS Journal of Photogrammetry and Remote Sensing, 103*, 48–56. https://doi.org/10.1016/j.isprsjprs.2014.06.016.

Seto, K. C., Sánchez-Rodríguez, R., & Fragkias, M. (2010). The new geography of contemporary urbanization and the environment. *Annual Review of Environment and Resources, 35*(1), 167–194. https://doi.org/10.1146/annurev-environ-100809-125336.

Smith, J. W., Smart, L. S., Dorning, M. A., Dupéy, L. N., Méley, A., & Meentemeyer, R. K. (2017). Bayesian methods to estimate urban growth potential. *Landscape and Urban Planning, 163*, 1–16. https://doi.org/10.1016/j.landurbplan.2017.03.004.

Stroud, J. R., Muller, P., & Sanso, B. (2001). Dynamic models for spatiotemporal data. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 63*(4), 673–689. https://doi.org/10.1111/1467-9868.00305.

Taubenbock, H., Esch, T., Felbier, A., Roth, A., & Dech, S. (2011). Pattern-based accuracy assessment of an urban footprint classification using TerraSAR-X data. *IEEE Geoscience and Remote Sensing Letters, 8*(2), 278–282. https://doi.org/10.1109/LGRS.2010.2069083.

Wu, J. (2014). Urban ecology and sustainability: The state-of-the-science and future directions. *Landscape and Urban Planning, 125*, 209–221. https://doi.org/10.1016/j.landurbplan.2014.01.018.

Zhang, C., Chen, Y., & Lu, D. (2015). Detecting fractional land-cover change in arid and semiarid urban landscapes with multitemporal Landsat Thematic mapper imagery. *GIScience & Remote Sensing, 52*(6), 700–722. https://doi.org/10.1080/15481603.2015.1071965.

Zhang, Q., & Seto, K. C. (2011). Mapping urbanization dynamics at regional and global scales using multi-temporal DMSP/OLS nighttime light data. *Remote Sensing of Environment, 115*(9), 2320–2329. https://doi.org/10.1016/j.rse.2011.04.032.

Zhou, Y., Smith, S. J., Zhao, K., Imhoff, M., Thomson, A., Bond-Lamberty, B., ... Elvidge, C. D. (2015). A global map of urban extent from nightlights. *Environmental Research Letters, 10*(5), 54011. https://doi.org/10.1088/1748-9326/10/5/054011.